International Journal of Hybrid Systems. **1** (2001), No 1, pp. 33-50.

A New Class of Hybrid Dynamical Systems: State Estimators with Bit-Rate Constraints

Nikolai G.Dokuchaev and Andrey V.Savkin

School of Electrical Engineering and Telecommunications, University of New South Wales,

Sydney, 2052, Australia

email: a.savkin@unsw.edu.au

ABSTRACT: This paper introduces a new class of hybrid dynamical systems. We consider a state estimation problem involving bit-rate communication capacity constraints for a discrete-time partially observed system. The observation must be coded and transmitted via a digital communication channel with a limited capacity. A recursive coder-estimator is proposed and investigated. An upper bound for the average estimation error is derived, and convergence properties are analyzed.
KEYWORDS: hybrid systems, state estimation, communication bit-rate constraints, Kalman filter
AMS (MOS) subject classification: 62M10, 93E11, 94A40

1. INTRODUCTION

In classical filtering theory (see e.g. Anderson and Moore [1]), the standard assumption is that all data transmission required by the algorithm can be performed with infinite precision. However, in some new models, it is common to encounter situations where observation and control signals are sent via a communication channel with a limited capacity. This problem may arise when a large number of mobile units need to be controlled remotely by a single decision maker. Since the radio spectrum is limited, communication constraint are a real concern. In Stitwell and Bishop [7], the problem of design of large-scale control systems for platoons of underwater vehicles highlights the need for control strategies that address reduced communications, since communication bandwidth is severely limited underwater. Another class of examples is offered by complex networked sensor systems containing a very large number of low power sensors. Furthermore, nowadays, it is becoming more common to use networks in systems, especially in those that are large-scale and physically distributed. All these new engineering applications motivate development of a new chapter of control and state estimation theory in which control and communication issues are combined together, and all the limitations of the communication channels are taken into account. Communications requirements, especially regarding bandwidth limits, are often challenging obstacles to control systems design. In these problems, classical Kalman estimation theory cannot be applied since the estimator only observes the transmitted sequence of finite-valued symbols. In fact, we need to design a hybrid dynamical system which consists of two subsystems. The first subsystem, that is called Coder, receives real-valued measurements and converts them into a finite-valued symbolic sequence which is sent over the limited capacity communication channel. The second subsystem (Decoder) receives this symbolic sequence and converts it into a real-valued state estimate. In other words, such state estimators with bit-rate constraints form an important subclass of so-called hybrid dynamical systems. In general, hybrid systems are those that combine continuous and discrete event dynamics and involve both real and symbolic variables; e.g., see Matveev and Savkin [3].

A natural question to ask is how much communication capacity is needed to achieve a specified estimation accuracy. The problem studied in this paper was introduced by Wong and Brockett [8], where some algorithms and models were proposed and investigated for the case of bounded random disturbances. A case of decreasing Gaussian disturbances was studied by Nair and Evans [4], where the idea to code the Kalman state estimate was proposed. However, the main results of these papers were restricted to the case of scalar systems.

In this paper, we investigate a state estimation problem involving constraints on bit-rate communication capacity for a discrete-time partially observed system of an arbitrary order with non-decreasing Gaussian disturbances. It is assumed that the observation must be coded and transmitted via a digital communication channel with a limited capacity. A recursive estimation algorithm is proposed and investigated for the case when system may be unstable. In this case, any large deviation of disturbances implies increasing all the following values of the state vector. We show that our algorithm provides state estimation with a bounded average error. Moreover, we obtain sufficient conditions of a convergence of the average error to zero as the digital communication channel capacity increases. As in the paper of Nair and Evans [4], our recursive coder-estimator includes the Kalman state estimator. It should be pointed out, that the proposed state estimation method is different from those described in literature; it is computationally non-expansive and easy to implement in real time. The most restrictive feature is that the algorithm is not adaptive to reducing of noise to zero, i.e. there is a given minimal level of tracking error which remains fixed even if the noise dissipates with time; we are not able to extend our proofs for a modification of the algorithm without this feature. (It can be added that we provided some numerical experiments with such modifications, but it appears that they lost stability; moreover, we have not found any proofs for other algorithms in literature for the case considered). However, the algorithm is adaptive to changing (i.e. increasing or decreasing) of the noise level *above some minimal level* which corresponds the minimal level of error. The obtained results can be extended to the case of uncertain linear systems (see e.g. Savkin and Petersen

[5], Savkin and Petersen [6]).

The remainder of this paper proceeds as follows. In Section 2, we introduce the class of systems under consideration and state the problem of estimation via limited capacity communication channels. Section 3 contains some well-known properties of the Kalman state estimator. In Section 4, we formulate the state estimation problem with communication constraints for a fully-observed system. Section 5 presents our recursive coding-estimation scheme. The main results of the paper are given in Section 6. Section 7 presents an illustrative example. Section 8 contains brief conclusions. The proofs of all the results of Section 6 are given in Appendix.

2. PROBLEM STATEMENT

Consider the following discrete-time linear system

$$\begin{cases} X_{t+1} = A_{t+1}X_t + B_tW_t, \\ Y_t = H_tX_t + D_tW_t, \end{cases} \tag{2.1}$$

where $X_t \in \mathbf{R}^n$ is the state, $W_t \in \mathbf{R}^d$ is the random disturbance input, $Y_t \in \mathbf{R}^m$ is the measured output, $t = 0, 1, 2, \ldots$.

We assume that the vectors X_0 and W_t are Gaussian, $\mathbf{E}|W_t|^2 \leq \text{const}\ \ (\forall t)$, $\mathbf{E}W_t \equiv 0$, and W_t does not depend on $W_0, W_1, ..., W_{t-1}$ and X_0.

Suppose estimates of the current state are required at a distant location, and are to be transmitted via a digital communication channel such that only M bits of data may be sent at each time t. We consider a system which consists of the coder, the transmission channel, and the decoder. Using an observation of $Y_1, ..., Y_t$, the coder produces a M-bit word h_t which is transmitted via the channel and then received by the decoder; the decoder produces an estimate $\widehat{X}_t$ which depends only on $h_1, ..., h_t$.

Let $\mathcal{A}$ be the set $\mathcal{A} = \{h\}$ of words $h = \left(h^{(1)}, ..., h^{(M)}\right)$, such that $h^{(i)} \in \{0, 1\}$. The set $\mathcal{A}$ consists of 2^M elements.

Let $h_t \in \mathcal{A}$ be the signal which is produced by the coder, $\widehat{X}_t$ be an estimate of X_t which is produced by the decoder.

Introduce the following vector and matrix norms:

$$\|x\| \triangleq \max_{i=1,...,n} |x^{(i)}| \quad \text{for} \quad x \in \mathbf{R}^n, \quad \|A\| \triangleq \max_{i=1,...,n} \sum_{j=1}^{n} |A^{(i,j)}| \quad \text{for} \quad A \in \mathbf{R}^{n\times n}.$$

We consider this problem as a problem of choosing the deterministic measurable functions $\Phi_t : (\mathbf{R}^m)^t \to \mathcal{A}$, and $F_t : (\mathcal{A})^t \to \mathbf{R}^n$, $t = 1, 2, ...$ such that

$$h_t = \Phi_t(Y_1, Y_2, ..., Y_t) \in \mathcal{A},$$

$$\widehat{X}_t = F_t(h_1, h_2, ..., h_t) \in \mathbf{R}^n,$$

and the following estimate holds:

$$\mathbf{E}|\widehat{X}_t - X_t| \leq \text{const} \quad (\forall t > 0).$$

Here $|\cdot|$ denotes the standard Euclidean norm. The main difficulty of this estimation problem is in a case of non-stable system, when any large deviation of disturbances implies increasing all the following values $|X_t|$.

It is well known that under some standard assumptions on A_t, B_t, H_t, D_t, there exists so-called Kalman estimate X_t^{KE} of X_t which minimizes the average error $\mathbf{E}|X_T^{KE} - X_T|^2$.

In this paper we propose an estimation algorithm which involves the Kalman estimation. The Kalman estimate is supposed to be computed, coded, transmitted via the channel and then decoded. The block diagram of our state estimation system is shown in Figure 2.1.

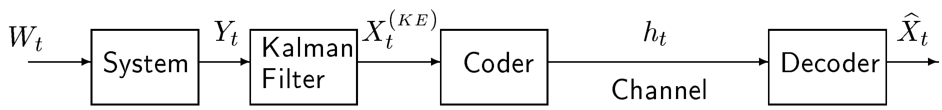

Figure 2.1: Block diagram of the estimator.

3. BASIC PROPERTIES OF THE KALMAN ESTIMATE

It is well known that the Kalman estimate X_t^{KE} satisfies the following equations:

$$\begin{cases} X_t^{KE} = A_t X_{t-1}^{KE} + V_t, \\ X_0^{KE} = \mathbf{E}X_0, \end{cases} \tag{3.1}$$

where

$$V_{t+1} = P_t\left(Y_t - H_t X_t^{KE}\right).$$

The matrix P_t is calculated recursively from the corresponding Riccati equation (see, e.g., Anderson and Moore [1]) and is uniformly bounded under some standard assumptions on the system. The estimation error

$$\Delta_t \triangleq X_t - X_t^{KE}$$

is independent on X_t^{KE}. The vectors X_t^{KE}, Δ_t are Gaussian, $\mathbf{E}\Delta_t \equiv 0$, $\mathbf{E}|\Delta_t|^2 \leq$ const (see, e.g., Anderson and Moore [1]).

Let us discuss basic properties of V_t. We have that

$$V_{t+1} = P_t\left(Y_t + H_t\Delta_t - H_t X_t\right) = P_t\left(D_t W_t + H_t \Delta_t\right).$$

Hence the vectors V_{t+1} are Gaussian, $\mathbf{E}V_t \equiv 0$, $\mathbf{E}|V_t|^2 \leq$ const $(\forall t > 0)$.

Furthermore,

$$\begin{aligned}\Delta_{t+1} &= A_{t+1}\Delta_t + B_t W_t - V_{t+1} \\ &= A_{t+1}\Delta_t + B_t W_t + P_t H_t X_t^{KE} - P_t Y_t \\ &= A_{t+1}\Delta_t + B_t W_t + P_t H_t \left(X_t - \Delta_t\right) - P_t Y_t.\end{aligned}$$

Hence

$$\Delta_{t+1} = \left(A_{t+1} - P_t H_t\right)\Delta_t + B_t W_t - P_t D W_t.$$

By the assumptions on W_t, we have that $\mathbf{E}(X_s^{KE})'\Delta_t = \mathbf{E}(X_s^{KE})'\Delta_s = 0$ for $s \le t$. Hence X_s^{KE} is independent on Δ_t, and $\mathbf{E}(X_s^{KE})'P_t H_t \Delta_t = 0$. Moreover,

$$\begin{aligned}&\mathbf{E}(X_s^{KE})'P_t D_t W_t = 0, \\ &\mathbf{E}(X_s^{KE})'V_{t+1} = \mathbf{E}(X_s^{KE})'P_t H_t \Delta_t + \mathbf{E}(X_s^{KE})'P_t D_t W_t = 0.\end{aligned}$$

Then the vector V_t is independent on $X_1^{KE}, X_2^{KE}, ..., X_{t-1}^{KE}$. As stated earlier, the vectors V_{t+1} are Gaussian, $\mathbf{E}V_t \equiv 0$, $\mathbf{E}|V_t|^2 \le \text{const}\ \ (\forall t > 0)$.

It may be concluded that the initial problem of estimation X_t may be stated as follows: *Estimate the state of the fully-observed system (3.1) under bit-rate constraints in the case of Gaussian disturbances V_t which do not depend on the previous states and have bounded variance.*

4. STATE ESTIMATION FOR THE FULLY-OBSERVED SYSTEM

Consider the process

$$\begin{cases} x_t = A_t x_{t-1} + b_t + v_t, \\ x_0 = \overline{x}_0. \end{cases} \tag{4.1}$$

Here $\overline{x}_0$ is a deterministic vector, v_t are random disturbances, $t \ge 0$, $x_t, v_t, b_t \in \mathbf{R}^n$ and $A_t \in \mathbf{R}^{n\times n}$. We assume that A_t, b_t and $\overline{x}_0$ are known.

In this paper, $\lfloor \alpha \rfloor$ denotes the integer part of a real number $\alpha > 0$, such that $\lfloor \alpha \rfloor = \max\{z \in \mathbf{Z} : z \le \alpha\}$.

Let

$$\overline{\theta} \triangleq 2^{M-1}, \quad \nu \triangleq \left\lfloor \overline{\theta}^{1/n} \right\rfloor, \qquad \theta \triangleq \nu^n, \quad L \triangleq \sup_{t\ge 0} \|A_t\|.$$

We suppose that the following assumptions hold:

Assumption 4.1 $\nu \ge 1$, $\nu > L$.

Assumption 4.2 *Vectors v_t are Gaussian, v_t is independent on $x_1, x_2, ..., x_{t-1}$,*

$$\mathbf{E}v_t = 0, \quad \mathbf{E}\|v_t\|^2 \le \delta^2 \quad (\forall t \ge 0),$$

where $\delta > 0$ is a given constant.

Also, we assume without a loss of generality that $L > 1$.

5. ESTIMATION ALGORITHM

The estimate $\widehat{x}_t$ of the process x_t will be found as a solution of the following equations:

$$\begin{cases} \widehat{x}_t = A_t \widehat{x}_{t-1} + b_t + c_t, \quad t > 0, \\ \widehat{x}_0 = \overline{x}_0, \end{cases} \tag{5.1}$$

Here

$$c_t = C_t(h_1, ..., h_t) \in \mathbf{R}^n, \quad h_t = \Phi_t(x_1, ..., x_t) \in \mathcal{A}. \tag{5.2}$$

The words h_t are to be calculated by the coder. The vectors c_t are to be calculated by the decoder. In (5.2), $\mathcal{A}$ is the set of words introduced in Section 2, $C_t : \mathcal{A}^t \to \mathbf{R}^n$, $\Phi_t : \mathbf{R}^{nt} \to \mathcal{A}$ are deterministic measurable functions. We consider the problem as a problem of choosing the functions $C_t(\cdot)$, $\Phi_t(\cdot)$ such that $\mathbf{E}\|x_t - \widehat{x}_t\|^r \le \text{const} \ (\forall t > 0)$ for a given constant $r \ge 1$.

We assume below that the set $\mathcal{A}$ is the set of pairs $\mathcal{A} = \{(\gamma, s)\}$, where $\gamma = 0$ or $\gamma = 1$, $s \in \{1, ..., \bar{\theta}\}$. Note that the set $\mathcal{A}$ consists of $2\overline{\theta} = 2^M$ elements and $\theta \le \overline{\theta}$.

Furthermore, let numbers $l > 0$, $a > 1$ and an integer $R \ge 1$ be given parameters.

For any $\lambda > 0$, set $D(\lambda) \triangleq \{x \in \mathbf{R}^n : \quad \|x\| \le \lambda\}$.

Consider a discrete subset $\widetilde{D}(\lambda) = \{y_j(\lambda)\}_{j=1}^{\theta} \subset D(\lambda)$ such that for any $x \in D(\lambda)$ there exists a vector $y \in \widetilde{D}(\lambda)$ such that $\|x - y\| \le \lambda\nu^{-1}$. It can be easily seen that such a subset $\widetilde{D}(\lambda)$ does exist.

For any $\lambda > 0$, introduce the following maps $S_1(x, \lambda) : \mathbf{R}^n \to \{1, ..., \theta\}$, $S_2(x, \lambda) : \mathbf{R}^n \to \{1, ..., \overline{\theta}\}$ and $F(x, \lambda) : \mathbf{R}^n \to \widetilde{D}(\lambda)$:

$$\begin{aligned} S_1(x, \lambda) &= \min\left\{ \arg\min{}_{j \in \{1, ..., \theta\}} \|y_j(\lambda) - x\| \right\}, \\ S_2(x, \lambda) &= \max\left\{ k \in \{1, ..., \overline{\theta}\} : \ x \notin D(\lambda a^{R(k-1)}) \right\}, \\ F(x, \lambda) &= y_j(\lambda), \quad \text{where} \quad j = S_1(x, \lambda). \end{aligned}$$

Note that if $x \in D(\lambda)$ and $S(x, \lambda) = j$ then $\|y_j(\lambda) - x\| \le \|y_i(\lambda) - x\| \ (\forall i = 1, ..., \theta)$ and $\|y_j(\lambda) - x\| \le \lambda\nu^{-1}$. If $x \in D(\lambda)$ then $\|F(x, \lambda) - x\| \le \lambda\nu^{-1}$.

Introduce the following vectors:

$$\widetilde{z}_t \triangleq x_t - A_t \widehat{x}_{t-1} - b_t, \quad t \ge 1. \tag{5.3}$$

Let $l_0 \triangleq l$. Then the following sequence of h_t, l_t, c_t is to be computed:

(i) The coder produces a word $h_t = (\gamma_t, s_t)$ and a number l_t, where

$$\gamma_t = \begin{cases} 0 & \text{if} \quad \widetilde{z}_t \in D(l_{t-1}) \\ 1 & \text{if} \quad \widetilde{z}_t \notin D(l_{t-1}) \end{cases}; \quad s_t = \begin{cases} S_1(\widetilde{z}_t, l_{t-1}), & \text{if } \gamma_t = 0 \\ S_2(\widetilde{z}_t, l_{t-1}), & \text{if } \gamma_t = 1, \end{cases}$$

$$l_t = \begin{cases} l_{t-1}/a & \text{if} \quad \gamma_t = 0,\ l_{t-1} > l/a \\ l_{t-1} & \text{if} \quad \gamma_t = 0,\ l_{t-1} = l/a \\ l_t = a^{Rs_t} l_{t-1} & \text{if} \quad \gamma_t = 1. \end{cases} \tag{5.4}$$

(ii) The word h_t is transmitted via the channel.

(iii) The decoder computes l_t by the rule (5.4), and then it calculates

$$c_t = \begin{cases} F(h_t, l_{t-1}) & \text{if} \quad \gamma_t = 0 \\ 0 & \text{if} \quad \gamma_t = 1. \end{cases} \tag{5.5}$$

(iv) Finally, the decoder computes $\widehat{x}_t$ by the formula (5.1).

6. THE MAIN RESULTS

In this section we show how to choose parameters l, a, R of the state estimation algorithm from Section 5 to guarantee that the average estimation error is bounded or converges to zero.

Introduce the process of the estimation error

$$z_t \triangleq x_t - \widehat{x}_t.$$

Theorem 6.1 *Consider the system (4.1) and the estimation algorithm described in the Section 5 with given parameters $a > 1$ and $R \geq 1$. Suppose that*

$$L \in \left(\frac{\nu}{a}, a^{R\overline{\theta}} \right). \tag{6.1}$$

Then for any $r \geq 1$ there exists a parameter $l > 0$ and a constant $C_* > 0$ such that

$$\mathbf{E}\|z_T\|^r \leq C_* \quad \forall T > 0. \tag{6.2}$$

Starting from now, we assume that $r > 1$ is fixed. To formulate sufficient conditions on l to guarantee (6.2) we need to introduce the following constants: Let $r \geq 1$ be a given number, and

$$\mu \triangleq \min\{i \in \{2, 4, 6, ..\} : i > r\}, \tag{6.3}$$

$$h \triangleq \frac{\nu - La}{\nu}, \quad G_i \triangleq \frac{\sqrt{2} n (i/2)! 2^{i/2}}{\sqrt{\pi}} \quad \text{for} \quad i = 2, 4, \ldots\ . \tag{6.4}$$

Theorem 6.2 *Suppose the assumptions (6.1) hold and the parameter $l > 0$ satisfies the inequality*

$$l^\mu \geq \frac{\delta^\mu G_\mu}{h^\mu (1 - a^{-r})(a^{\mu - r} - 1)}. \tag{6.5}$$

Then there exists a constant C_* such that the inequality (6.2) holds.

To give an upper estimate of C_* we need to introduce the following constants:

$$\beta \triangleq a^{R\overline{\theta}}, \quad C_0 \triangleq \frac{\delta^2}{(L-1)^2}, \tag{6.6}$$

Furthermore, for any $\rho \geq 1$, introduce the following constants:

$$q_\rho \triangleq \min\left\{ i \in \{2,4,6,..\} : \quad i > \rho, \quad \frac{\beta^\rho L^i}{\beta^i} < 1 \right\}, \quad Q_\rho \triangleq \frac{\beta^\rho L^{q_\rho}}{\beta^{q_\rho}}, \tag{6.7}$$

$$\overline{\delta}(\rho) \triangleq \sup_{t \geq 0} \left(\mathbf{E}\|v_t\|^\rho\right)^{1/\rho}, \quad f(\rho) \triangleq \beta^\rho \left(1 + \frac{G_{q_\rho} \beta^{2q_\rho+\rho} C_0^{q_\rho/2}}{l^{q_\rho}(\beta - L)^{q_\rho}} \frac{Q_\rho}{1 - Q_\rho}\right), \tag{6.8}$$

Theorem 6.3 *Under the assumptions of Theorem 6.2, the inequality (6.2) holds with*

$$C_* = \frac{l^r}{\nu^r} + \left(3l f(r)^{1/r} + 2f(r\rho')^{1/r\rho'} \frac{\overline{\delta}(r\rho)}{L-1}\right)^r, \tag{6.9}$$

where $\rho > 1$ is an arbitrary number, $\rho' \triangleq \rho(\rho-1)^{-1}$.

The following theorem gives sufficient condition for a convergence for the case of increasing channel capacity M.

Theorem 6.4 *Suppose a time $T > 0$ is fixed, and $l^{-1} + l\nu^{-1} \to 0$, where $\nu = \left\lfloor 2^{\frac{M-1}{n}} \right\rfloor$. Then $\mathbf{E}\|z_t\|^r \to 0$ uniformly on $t \leq T$.*

The proofs of all the results of this section are given in Appendix.

7. ILLUSTRATIVE EXAMPLE

To illustrate the results of this paper, we consider a deconvolution problem similar to those considered in Chen and Chen [2]. The block diagram of deconvolution system is shown in Figure 7.1.

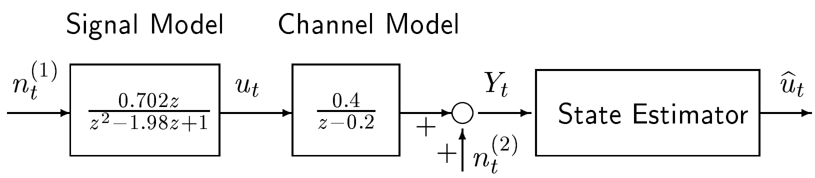

Figure 7.1: Deconvolution system.

Combining the signal model and the channel model, we obtain the following system:

$$\begin{aligned}
\begin{bmatrix} X_{t+1}^{(1)} \\ X_{t+1}^{(2)} \\ X_{t+1}^{(3)} \end{bmatrix} &= \begin{bmatrix} 1.98 & -1 & 0 \\ 1 & 0 & 0 \\ 0.4 & 0 & 0.2 \end{bmatrix} \begin{bmatrix} X_t^{(1)} \\ X_t^{(2)} \\ X_t^{(3)} \end{bmatrix} + \begin{bmatrix} 0.707 \\ 0 \\ 0 \end{bmatrix} n_t^{(1)}, \\
Y_t &= X_t^{(3)} + n_t^{(1)}.
\end{aligned}$$

In this state space description, $X_t^{(1)}$ and $X_t^{(2)}$ are the state variables of the signal model. In addition, $u_t = X_t^{(3)}$ is the state variable of the channel model, Y_t is the measured signal. To apply our results to this deconvolution problem, we consider a corresponding system of the form (2.1). In this case, the matrices A_t, B_t, K_t, D_t are given by

$$A = \begin{bmatrix} 1.98 & -1 & 0 \\ 1 & 0 & 0 \\ 0.4 & 0 & 0.2 \end{bmatrix}, \quad B_t = \begin{bmatrix} 0.707 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$K_t = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, \quad D_t = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

Also, the constants n, d, L are given by $n = 3$, $d = 2$, $L = 2.98$.

To illustrate the performance of our coder-estimator, we consider the Gaussian noise signal $W_t = (n_t^{(1)}, n_t^{(2)})$ with $\mathbf{E}W_t \equiv 0$, $\mathbf{E}W_tW_t' \equiv 0.8 \cdot I$. Also, we take initial condition $X_0 = (1, -1, 1)'$. We consider the system in the cases of a communication channel with capacity $M = 8$ and $M = 10$ bits. We apply the estimation algorithm from Section 5 with the parameters $l = 10.2$, $R = 20$, and $a = 1.3$ for $M = 8$, $a = 1.53$ for $M = 10$. Figure 7.2 shows the true value u_t, the Kalman estimate, and the resulting estimates of the signal u_t for times $t = \overline{150, 200}$ in the cases of communication channel with capacity $M = 8$ and $M = 10$ bits. Figure 7.3 shows the true value u_t and the resulting estimate of the signal u_t for $t = \overline{1, 200}$ in the case of the communication channel with capacity $M = 10$ bits.

8. CONCLUSIONS

This paper describes a new class of hybrid dynamical systems. It considers a state estimation problem involving bit-rate communication capacity constraints for a discrete-time partially observed system. The observation must be coded and transmitted via a digital communication channel with a limited capacity. Classical estimation theory cannot be applied since the estimator only observes the transmitted sequence of finite-valued symbols. A recursive estimation algorithm is proposed and investigated. We show that our algorithm provides state estimation with a bounded average error which converges to zero as the digital communication channel capacity increases. The proposed state estimation method is computationally non-expansive and easy to implement in real-time systems.

9. ACKNOWLEDGEMENT

This work was supported by the Australian Research Council.

REFERENCES

[1] B.D.O. Anderson, J.B. Moore, Optimal Filtering, Englewood Cliffs, NJ: Prentice-Hall, 1979.

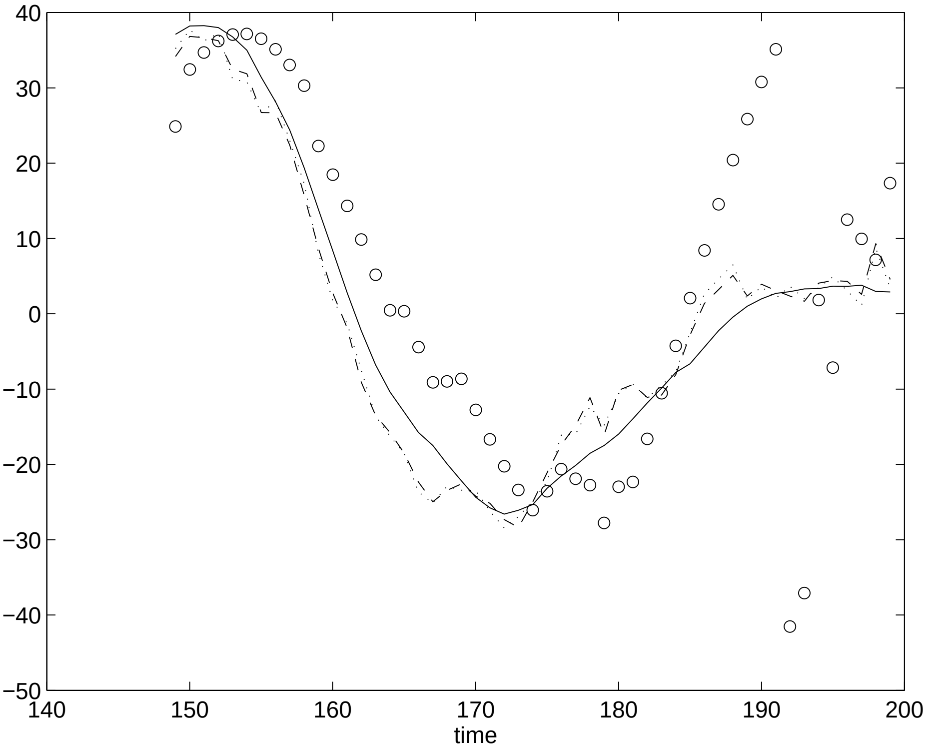

Figure 7.2: Estimates of u_t for $t = \overline{150, 200}$; "—" - true value of u_t; "$- - -$" - Kalman estimate; "ooo" - estimate for 8 bit channel; "..." - estimate for 10 bit channel

[2] Y. Chen, B.S. Chen, Minimax robust deconvolution filters under stochastic parameters and noise uncertainties. IEEE Trans. Signal Processing, **42** (1994) 32-45.

[3] A.S. Matveev, A.V. Savkin, Qualitative Theory of Hybrid Dynamical Systems. Birkhauser, Boston, 2000.

[4] G. Nair, R.J. Evans, State estimation under bit-rate constraints. In Proc. 37th IEEE Conf. Decision and Control, 1998.

[5] A.V. Savkin, I.R. Petersen, Recursive state estimation for uncertain systems with an integral quadratic constraint. IEEE Trans. Automat. Contr., **40**(6) (1995) 1080-1083.

[6] A.V. Savkin, I.R. Petersen, Robust state estimation and model validation for discrete-time uncertain systems with deterministic description of noise and uncertainty, Automatica, **34**(2) (1998) 271-274.

[7] D.J. Stilwell, B.E. Bishop, Platoons of underwater vehicles, IEEE Control Systems Magazine., **20**(6) (2000) 45-52.

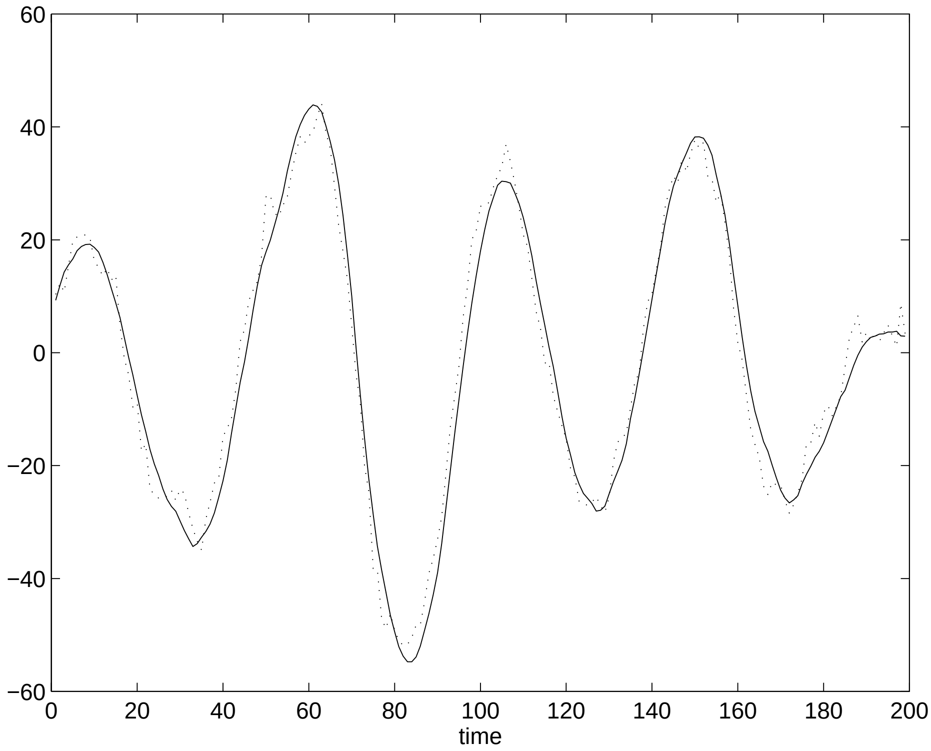

Figure 7.3: Estimates of u_t for $t = \overline{1,200}$; "—" - true value of u_t; "..." - estimate for 10 bit channel

[8] W.S. Wong, R.W. Brockett, Systems with finite communication bandwidth constraints - Part 1: state estimation problems. *IEEE Trans. Automat. Contr.*, **42**(9), (1997) 1294-1299.

APPENDIX: PROOFS

Let l_t, c_t and z_t be computed by the algorithm, $t = 1, 2, \ldots$.

Introduce a sequence $\{s_i, \tau_i, t_i\}_{i=0}^{+\infty}$ of triplets of integer random times, such that the following conditions hold:

(i) $s_0 = \tau_0 = 0$, $1 \le s_i \le \tau_i < t_i < s_{i+1}$ $(\forall i \ge 1)$;

(ii) if $s_1 = 1$ then $t_0 = 0$; if $s_1 > 1$ then $t_0 = 1$;

(iii) if $t \notin \cup_{i \ge 1} \{s_i, s_i + 1, ..., t_i - 1, t_i\}$ and $t > 0$ then $l_t = l/a$;

(iv) if $t \in \cup_{i \ge 1} \{s_i, s_i + 1, ..., \tau_i - 1, \tau_i\}$ then $l_t > l/a$, $\gamma_t = 1$;

(v) if $t \in \cup_{i \ge 1} \{\tau_i + 1, ..., t_i - 1\}$ then $l_t > l/a$, $\gamma_t = 0$;

(vi) if $t \in \cup_{i \ge 1} \{t_i : \ t_i < s_{i+1} - 1\}$ then $l_t = l/a$, $\gamma_t = 0$.

Introduce the following random sequences:

$$\xi_i \triangleq \tau_i - s_i + 1, \quad \eta_i \triangleq t_i - \tau_i \quad \text{for} \quad i = 0, 1, 2, \dots ,$$

$$\zeta_0 \triangleq -\eta_0, \quad \zeta_i \triangleq \zeta_{i-1} + \log_a \frac{l_{\tau_i}}{l_{s_i - 1}}, \quad \text{for} \quad i = 1, 2, \dots . \tag{A.1}$$

We assume that

$$z_{-1} = 0, \quad s_{-1} = \tau_{-1} = t_{-1} = 0, \quad \xi_{-1} = 0, \quad \eta_{-1} = 0, \quad \zeta_{-1} = 0.$$

Let $T > 0$ be a fixed deterministic integer number. Introduce random integer variables m and k such that $s_m \le T < s_{m+1}$ and $k = m - 1$. Let $\mathcal{F}'$ be the σ-algebra of random events which is generated by the random values $\{\zeta_k, \tau_k\}$, and $\mathcal{F}''$ be the σ-algebra of random events which is generated by the random values $\{\zeta_k, \tau_k, \eta_m, s_m, z_{s_m - 1}\}$.

We will use the notation Ind for the indicator function.

Proposition A.1 *If $\gamma_t = 0$ then $\|z_t\| \le \frac{l_{t-1}}{\nu} \le l_t$.*

Proof. The equation (5.3) can be rewritten as follows:

$$\widetilde{z}_t = A_t x_{t-1} - A_t \widehat{x}_{t-1} + v_t. \tag{A.2}$$

Hence

$$\widetilde{z}_t = A_t z_{t-1} + v_t. \tag{A.3}$$

We have that $z_t = \widetilde{z}_t - c_t$,

$$z_t = A_t z_{t-1} + v_t - c_t, \quad z_0 = 0. \tag{A.4}$$

Then $\|z_t\| = \| \le l_{t-1}/\nu$, and $\nu > a$. Hence $\|z_t\| \le l_t$. This completes the proof.□

Proposition A.2 *If $\tau_j \le t < s_{j+1}$, $j \ge 1$, then $l_t \le l a^{\zeta_j - \eta_j \wedge (t - \tau_j)}$. If $s_j \le t \le \tau_j$, $j \ge 1$, then*

$$l a^{\zeta_{j-1} - \eta_{j-1} + R(\xi_j \wedge (t - s_j + 1))} \le l_t \le l a^{\zeta_{j-1} - \eta_{j-1} + R\overline{\theta}(\xi_j \wedge (t - s_j + 1))}, \quad j = 1, 2, \dots .$$

Proof. Proposition A.2 follows immediately from the description of the algorithm.□

Proposition A.3 *The random variable* $\mathrm{Ind}\{k \ge 1\}$ *is $\mathcal{F}'$–measurable and $\mathcal{F}''$–measurable.*

Proof. We have that $k < 1$ if and only if $\zeta_k \le 0$, $\tau_k = 0$. It means that $\mathrm{Ind}\{k \ge 1\}$ is $\mathcal{F}'$–measurable. Furthermore, $\mathcal{F}' \subset \mathcal{F}''$, hence $\mathrm{Ind}\{k \ge 1\}$ is $\mathcal{F}''$–measurable. This completes the proof of Proposition.

Proposition A.4 *Let* $\psi = \left(\psi^{(1)}, \psi^{(2)}, ..., \psi^{(n)}\right)$ *be a Gaussian random n–dimensional vector such that*

$$|\mathbf{E}\psi^{(i)}| \le \alpha, \quad \mathrm{Var}\,\psi^{(i)} \le c^2 \quad \forall i = 1, ..., n,$$

where $\alpha \ge 0$, $c > 0$ *are fixed. Then the following estimate holds:*

$$\mathbf{P}(\|\psi\| > u) \le \frac{G_\kappa c^\kappa}{|u-\alpha|^\kappa} \quad \forall u > \alpha, \quad \forall \kappa = 2, 4, 6, ..,$$

where the constants G_κ *are defined by (6.4).*

Proof. (i) Let $n = 1$, $\mathbf{E}\psi = 0$, $\mathbf{E}\psi^2 = 1$. Then

$$\begin{aligned}\mathbf{P}\,(\psi > u) \; &= \mathbf{P}\,(\psi < -u) = \tfrac{1}{\sqrt{2\pi}} \textstyle\int_u^{+\infty} t e^{-\frac{t^2}{2}} dt \\ &= \tfrac{1}{\sqrt{2\pi}} \textstyle\int_{u^2/2}^{+\infty} e^{-y} dy = \tfrac{1}{\sqrt{2\pi}} e^{-u^2/2} = \tfrac{1}{\sqrt{2\pi}} \left(\sum_{i=0}^{+\infty} \tfrac{1}{i!} \left(\tfrac{u^2}{2}\right)^i \right)^{-1} \le \tfrac{1}{\sqrt{2\pi}} \tfrac{q! 2^q}{u^{2q}}\end{aligned}$$

for all $q = 1, 2, ...$. This completes the proof of Proposition for this case.

(ii) Let $n = 1$, $\mathbf{E}\psi = 0$, $\mathbf{E}\psi^2 = c^2$. Let $\psi / c = \psi_c$. Then

$$\mathbf{P}\,(\psi > u) = \mathbf{P}\,(\psi < -u) = \mathbf{P}\left(\psi_c > \frac{u}{c}\right) = \mathbf{P}\left(\psi_c < -\frac{u}{c}\right) \le \frac{1}{\sqrt{2\pi}} \frac{q! 2^q c^{2q}}{u^{2q}}$$

for all $q = 1, 2, ...$. This completes the proof of Proposition for this case.

(iii) Let $n = 1$, $\mathbf{E}\psi = \overline{\psi}$, $|\overline{\psi}| \le \alpha$. Let $\psi_0 = \psi - \overline{\psi}$, $\mathbf{E}\psi_0^2 = c^2$. Then $\mathbf{E}\psi_0 = 0$, and

$$\begin{aligned}\mathbf{P}\,(|\psi| > u) \; &= \mathbf{P}\left(\psi_0 < -u - \overline{\psi}\right) + \mathbf{P}\left(\psi_0 > u - \overline{\psi}\right) \\ &\le \tfrac{1}{\sqrt{2\pi}} \tfrac{q! 2^q c^{2q}}{(u - \overline{\psi})^{2q}} + \tfrac{1}{\sqrt{2\pi}} \tfrac{q! 2^q c^{2q}}{(u + \overline{\psi})^{2q}} \le \tfrac{\sqrt{2}}{\sqrt{\pi}} \tfrac{q! 2^q c^{2q}}{(u-\alpha)^{2q}}\end{aligned}$$

for all $q = 1, 2, ...$. This completes the proof of Proposition for this case.

(iv) Let $n > 1$, then

$$\mathbf{P}(\|\psi\| > u) = \mathbf{P}\left(\exists i : \left|\psi^{(i)}\right| > u\right) \le n \max_{i=1,...,n} \mathbf{P}(\|\psi^{(i)}\| > u) \le \frac{\sqrt{2}}{\sqrt{\pi}} \frac{n\, q! 2^q c^{2q}}{(u-\alpha)^{2q}}$$

for all $q = 1, 2, ...$. This completes the proof of Proposition. □

We assume below that a number $r \ge 1$ is fixed, and the number μ is defined in (6.3). We use the notations q, Q for the constants q_r, Q_r defined in (6.7) with $\rho = r$.

Lemma A.1 *The following estimate holds:*

$$\mathrm{Ind}\,\{k \ge 1\} \left(a^{r\zeta_k} \mathbf{E}\,\{a^{-r\eta_k} | \mathcal{F}'\} - 1 \right) \le 0 \quad a.s.$$

Proof. For $j \ge 0$, introduce random events

$$\Omega^{(j)} \triangleq \{\|\widetilde{z}_{\tau_k+j}\| \le l_{\tau_k+j-1}, \quad \|\widetilde{z}_{\tau_k+j+1}\| > l_{\tau_k+i}\}\,. \tag{A.5}$$

Substituting (5.5) into (A.4), we have that the event $\Omega^{(j)}$ implies

$$\|z_{\tau_k+j}\| \le \frac{l_{\tau_k+j-1}}{\nu}, \quad \|A_{\tau_k+j+1} z_{\tau_k+j} + v_{\tau_k+j+1}\| \ge l_{\tau_k+j}.$$

Hence
$$\|A_{\tau_k+j+1}z_{\tau_k+j}\| \le \frac{Ll_{\tau_k+j-1}}{\nu} \le \frac{aLl_{\tau_k+j}}{\nu}.$$
It follows that the event $\Omega^{(j)}$ implies
$$\|v_{\tau_k+j+1}\| \ge l_{\tau_k+j}\left(1-\frac{La}{\nu}\right) = l_{\tau_k+j}h. \tag{A.6}$$
Furthermore,
$$\mathbf{P}\left(\eta_k = j\,|\mathcal{F}'\right) \le \mathbf{P}\left(\|\widetilde{z}_{\tau_k+j+1}\| > l_{\tau_k+i},\quad \|\widetilde{z}_{\tau_k+j}\| \le l_{\tau_k+j-1}|\mathcal{F}'\right), \quad j\ge 1.$$
Let $\omega \triangleq G_\mu\delta^\mu h^{-\mu}$. We have that
$$\mathbf{P}\left(\eta_k = j\,|\mathcal{F}'\right) \le \mathbf{P}\left(\|v_{\tau_k+j+1}\| \ge l_{\tau_k+j}h|\mathcal{F}'\right).$$
It follows from Proposition A.4 that
$$\mathbf{P}\left(\eta_k = j\,|\mathcal{F}'\right) \le \frac{G_\mu\delta^\mu}{l^\mu_{\tau_k+j}h^\mu} = \frac{\omega}{l^\mu_{\tau_k+j}}.$$
By (A.1), we have that $l_{\tau_k+j} = a^j l_{\tau_k}$. Moreover, from Proposition A.2 we have that $l_{\tau_k} = a^{\zeta_k} l a^{-1}$. Hence
$$\mathbf{P}\left(\eta_k = j\,|\mathcal{F}'\right) \le \frac{\omega a^{\mu j}}{l^\mu_{\tau_k}} = \frac{\omega}{l^\mu}a^{\mu j-\mu\zeta_k}.$$
Then
$$\sum_{j=1}^{\zeta_k} a^{-rj}\mathbf{P}\left(\eta_k = j|\mathcal{F}'\right) \le \tfrac{\omega}{l^\mu}\sum_{j=1}^{\zeta_k} a^{\mu j-\mu\zeta_k-rj}$$
$$\le \tfrac{\omega}{l^\mu}a^{-\mu\zeta_k}\sum_{j=1}^{\zeta_k} a^{j(\mu-r)} \le \tfrac{\omega}{l^\mu}a^{-\mu\zeta_k}a^{\mu-r}\tfrac{a^{(\zeta_k-1)(\mu-r)}-1}{a^{\mu-r}-1} \le \tfrac{\omega}{l^\mu}\tfrac{a^{-\mu\zeta_k+\mu-r+(\zeta_k-1)(\mu-r)}}{a^{\mu-r}-1} \le \tfrac{\omega}{l^\mu}\tfrac{a^{-r\zeta_k}}{a^{\mu-r}-1}.$$
Hence
$$\begin{aligned}&\mathrm{Ind}\,\{k\ge 1\}\mathbf{E}\left\{a^{-r\eta_k}|\mathcal{F}'\right\}\\ &\le \mathrm{Ind}\,\{k\ge 1\}\left[\textstyle\sum_{j=1}^{\zeta_k} a^{-rj}\mathbf{P}\left(\eta_k = j\,|\mathcal{F}'\right) + a^{-r\zeta_k-r}\mathbf{P}\left(\eta_k \ge \zeta_k+1|\mathcal{F}'\right)\right]\\ &\le \mathrm{Ind}\,\{k\ge 1\}\left(\tfrac{\omega}{l^\mu}\tfrac{a^{-r\zeta_k}}{a^{\mu-r}-1} + a^{-r\zeta_k-r}\right) \le \mathrm{Ind}\,\{k\ge 1\}a^{-r\zeta_k}\left(\tfrac{\omega}{l^\mu}\tfrac{1}{a^{\mu-r}-1}+a^{-r}\right) \quad \text{a.s.}\end{aligned}$$
The equation (6.5) implies that
$$\frac{\omega}{l^\mu}\frac{1}{a^{\mu-r}-1} + a^{-r} \le 1.$$
This completes the proof of Lemma A.1.$\square$

Introduce the following matrices:
$$\overline{A}(t,s) \triangleq A_tA_{t-1}\cdots A_s \quad \text{for} \quad s\le t, \qquad \overline{A}(t-1,t)\triangleq I. \tag{A.7}$$
Here I is the unit matrix. Furthermore, introduce the following random vectors:
$$\overline{z}_{t,m} \triangleq \overline{A}(t,s_m)z_{s_m-1} + \sum_{j=s_m}^{t}\overline{A}(t,j+1)v_j, \quad \text{where} \quad t = s_m+1, s_m+2, \dots\,.$$

Proposition A.5 *For the conditional probabilistic measure $\mathbf{P}(\,\cdot\,|\mathcal{F}'')$, the random vectors $\overline{z}_{t,m}$ are Gaussian, and the following inequalities hold:*

$$\|\mathbf{E}\{\overline{z}_{t,m}|\mathcal{F}''\}\| \le L^{t-s_m+1}\frac{l_{s_m-1}}{\nu} \quad a.s., \tag{A.8}$$

$$\mathbf{E}\left\{\|\overline{z}_{t,m}-\mathbf{E}\{z_{t,m}|\mathcal{F}''\}\|^2\,|\mathcal{F}''\right\} \le C_0 L^{2(t-s_m+1)} \quad a.s., \tag{A.9}$$

where C_0 is defined by (6.6).

Proof. By Assumption 4.1, the vector v_{t+1} does not depend on $x_t, x_{t-1}, ..$ for any deterministic t. Moreover, (5.2) implies that there exist deterministic measurable functions $\phi_t : \mathbf{R}^{nt} \to \mathbf{R}^n$ such that $\widehat{x}_t = \phi_t(x_1, ..., x_t)$. We have that $z_t = x_t - \widehat{x}_t$. Hence v_{t+1} does not depend on $z_t, z_{t-1}, ...$. Furthermore, the vectors v_{s_m+i}, $i = 1, 2, ..$, are Gaussian and independent on z_{s_m-1} for the conditional probabilistic measure $\mathbf{P}(\,\cdot\,|\mathcal{F}'')$, and $\mathbf{E}\{\overline{z}_{t,m}|\mathcal{F}''\} = \overline{A}(t, s_m)z_{s_m-1}$.

Furthermore,

$$\|\widetilde{z}_{s_m-1}\| \le \frac{l_{s_m-1}}{\nu}, \quad \|\overline{A}(t,j)\| \le L^{t-j+1}.$$

Hence (A.8) holds. Further,

$$\begin{aligned}
&\mathbf{E}\left\{\|\overline{z}_{t,m}-\mathbf{E}\{\overline{z}_{t,m}|\mathcal{F}''\}\|^2\,\middle|\mathcal{F}''\right\} = \mathbf{E}\left\{\left(\textstyle\sum_{j=s_m}^{t}\overline{A}(t,j+1)v_j\right)^2|\mathcal{F}''\right\}\\
&\le \left[\textstyle\sum_{j=s_m}^{t}\left(\|\overline{A}(t,j+1)\|\mathbf{E}\{\|v_j\|^2|\mathcal{F}''\}\right)^{1/2}\right]^2\\
&\le \delta^2\left(\textstyle\sum_{j=s_m}^{t}L^{t-j}\right)^2 = \delta^2\left(\textstyle\sum_{j=0}^{t-s_m}L^{t-j-s_m}\right)^2 \le \delta^2\left(L^{t-s_m}\textstyle\sum_{j=0}^{t-s_m}L^{-j}\right)^2\\
&\le \delta^2\left(L^{t-s_m}\tfrac{1}{1-L^{-1}}\right)^2 \le \delta^2\tfrac{L^{2(t-s_m+1)}}{(L-1)^2} = C_0L^{2(t-s_m+1)}.
\end{aligned}$$

Hence (A.9) holds. This completes the proof of Proposition.

Proposition A.6 *The following estimate holds:*

$$\mathbf{P}(\xi_m \ge i+1|\mathcal{F}'') \le G_q\frac{L^{iq}\beta^{2q}C_0^{q/2}}{l^q\beta^{iq}(\beta-L\nu^{-1})^q} \quad \forall i\ge 1, \quad \forall q = 2,4,6,...$$

Proof. After substituting (5.5) into (A.4), we obtain that

$$\widetilde{c}_{s_m+i} = 0, \quad z_{s_m+i} = \overline{z}_{s_m+i,m} = \widetilde{z}_{s_m+i} \quad \text{for} \quad i = 0, ..., \tau_m - s_m. \tag{A.10}$$

Moreover, it can be easily seen, that

$$\mathbf{P}(\xi_m \ge i+1|\mathcal{F}'') \le \mathbf{P}(\|z_{s_m+i}\| > l_{s_m+i-1}|\mathcal{F}''), \quad i \ge 1.$$

Furthermore, if $s_m \le i \le \tau_m$ then

$$\begin{aligned}
&l_{s_m-1} \ge \tfrac{l}{a}, \quad l_{s_m+i-1} \ge l_{s_m-1}\beta^i,\\
&l_{s_m+i-1} - \tfrac{l_{s_m-1}}{\nu}L^i \ge l_{s_m-1}\left(\beta^i - \tfrac{L^i}{\nu}\right) \ge \tfrac{l\omega_1\beta^{(i-1)}}{a} \ge l\omega_1\beta^{i-2},
\end{aligned}$$

where

$$\omega_1 \triangleq \beta - L.$$

By Propositions A.4-A.5 and (A.10), it follows that

$$\mathbf{P}\left(\xi_m \geq i+1 | \mathcal{F}''\right) \leq \frac{G_q L^{iq} C_0^{q/2}}{(l_{s_m+i-1} - l_{s_m-1} L^i/\nu)^q} \leq \frac{G_q \beta^{2q} L^{iq} C_0^{q/2}}{\omega_1^q l^q \beta^{iq}} \quad \forall q = 2, 4, 6, ...$$

This completes the proof of Proposition A.6.

Lemma A.2 *The following estimate holds:*

$$\mathbf{E}\left\{\beta^{r\xi_m} | \mathcal{F}''\right\} \leq f(r) \quad a.s.,$$

where $f(\cdot)$ defined by (6.8).

Proof of Lemma A.2. We have that $\mathbf{P}\left(\xi_m \geq 1 | \mathcal{F}''\right) \leq 1$ and

$$\mathbf{E}\left\{\beta^{r\xi_m} | \mathcal{F}''\right\} \leq \beta^r + \sum_{i=1}^{+\infty} \beta^{(i+1)r} \mathbf{P}\left(\xi_m = i+1 | \mathcal{F}''\right) \leq \beta^r + \frac{G_q \beta^{2q+r} C_0^{q/2}}{l^q \omega_1^q} \sum_{i=1}^{+\infty} Q^i,$$

where $Q = Q_r$. Then

$$\mathbf{E}\left\{\beta^{r\xi_m} | \mathcal{F}''\right\} \leq \beta^r + \frac{G_q \beta^{2q+r} C_0^{q/2}}{l^q \omega_1^q} \frac{Q}{1-Q},$$

where $q = q_r$. This completes the proof of Lemma A.2.

Proof of Theorem 6.1–6.3. Let m and k be the random numbers defined above for the integer $T > 0$. Consider the following random events:

$$\Omega_0 \triangleq \{T < s_1\}, \quad \Omega_1 \triangleq \{s_1 \leq T \leq \tau_1\}, \quad \Omega_2 \triangleq \{\tau_1 < T < s_2\}, \quad \Omega_3 \triangleq \{s_2 \leq T\}. \tag{A.11}$$

It is easy to see that

$$\mathbf{P}\left(\cup_{i=0}^3 \Omega_i\right) = 1, \quad \mathbf{P}\left(\Omega_i \cap \Omega_j\right) = 0, \quad i \neq j, \quad i, j = \overline{0, 3},$$

$$\Omega_0 = \{m = 0\}, \quad \Omega_1 \cup \Omega_2 \cup \Omega_3 = \{m = 1\}, \quad \Omega_3 = \{m \geq 2\} = \{k \geq 1\}.$$

We have that

$$\mathbf{E}\|z_T\|^r \leq \mathbf{E}\left\{\|z_T\|^r | \Omega_0\right\} \mathbf{P}(\Omega_0) + \sum_{i=1}^3 \mathbf{E}\,\mathrm{Ind}\left\{\Omega_i\right\} \|z_T\|^r.$$

If $T < s_1$ then $\|z_T\| \leq l/\nu$, hence

$$\mathbf{E}\left\{\|z_T\|^r | \Omega_0\right\} \leq \frac{l^r}{\nu^r}. \tag{A.12}$$

Set

$$V \triangleq \sum_{j=s_m}^{T} \overline{A}(T, j+1) v_j.$$

If $s_m \le T \le \tau_m$, then $z_T = \overline{A}(T, s_m) z_{s_m - 1} + V$, where $\overline{A}(T, j)$ is defined by (A.7). By (6.1),

$$\begin{aligned} &\|\overline{A}(T, s_m)\| \le L^{T - s_m + 1} \le a^{R\xi_m \overline{\theta}}, \\ &\|V\| \le \textstyle\sum_{j=s_m}^T L^{T-j} \|v_j\| \le L^{T - s_m + 1} \sum_{j=s_m}^T L^{s_m - j - 1} \|v_j\|. \end{aligned} \tag{A.13}$$

Consider the random event $\Omega' \stackrel{\Delta}{=} \{s_m \le T \le \tau_m\}$. We have that

$$\begin{aligned} &\mathbf{E}\,\mathrm{Ind}\,\{\Omega'\}\|V\|^r \le \mathbf{E}\Big(\mathrm{Ind}\,\{\Omega'\} L^{T-s_m+1} \textstyle\sum_{j=s_m}^T L^{s_m - j - 1}\|v_j\|\Big)^r \\ &\le \Big(\mathbf{E}\,\mathrm{Ind}\,\{\Omega'\} L^{r\rho'(T-s_m)}\Big)^{1/\rho'} \Big(\mathbf{E}\Big(\textstyle\sum_{j=s_m}^T L^{s_m-j-1}\|v_j\|\Big)^{r\rho}\Big)^{1/\rho} \\ &\le \Big(\mathbf{E} L^{r\rho'\xi_m}\Big)^{1/\rho'} \Big(\textstyle\sum_{j=s_m}^T L^{s_m-j-1} \left(\mathbf{E}\|v_j\|^{r\rho}\right)^{1/r\rho}\Big)^r \le f(r\rho')^{1/\rho'} \frac{\overline{\delta}(r\rho)^r}{(L-1)^r}. \end{aligned} \tag{A.14}$$

Further, we have that $\|z_t\| \le l/\nu$ for $t \le s_1 - 1$ and $\Omega_1 \subset \Omega'$. Clearly,

$$\Big(\mathbf{E}\,\mathrm{Ind}\,\{\Omega_1\}\|z_T\|^r\Big)^{1/r} \le \Big(\mathbf{E}\,\mathrm{Ind}\,\{\Omega_1\}\|\overline{A}(T, s_1) z_{s_1-1}\|^r\Big)^{1/r} + \Big(\mathbf{E}\,\mathrm{Ind}\,\{\Omega_1\}\|V\|^r\Big)^{1/r}.$$

Then (A.13) and Lemma A.2 imply that

$$\begin{aligned} \Big(\mathbf{E}\,\mathrm{Ind}\,\{\Omega_1\}\|z_T\|^r\Big)^{1/r} &\le \tfrac{l}{\nu}\Big(\mathbf{E}\,\mathrm{Ind}\,\{\Omega_1\} L^{r\xi_1}\Big)^{1/r} + f(r\rho')^{1/r\rho'} \tfrac{\overline{\delta}(r\rho)}{L-1} \\ &\le \tfrac{l}{\nu} f(r)^{1/r} + f(r\rho')^{1/r\rho'} \tfrac{\overline{\delta}(r\rho)}{L-1}. \end{aligned} \tag{A.15}$$

If $\tau_1 < T \le s_2$, then $\|z_T\| \le l_t \le l\beta^{\xi_1}$, hence

$$\mathbf{E}\,\mathrm{Ind}\,\{\Omega_2\}\|z_T\|^r \le l^r \mathbf{E}\,\mathrm{Ind}\,\{\Omega_2\}\beta^{\xi_1} = l^r \mathbf{E}\,\mathrm{Ind}\,\{\Omega_2\}\beta^{\xi_m} \le l^r f(r). \tag{A.16}$$

By the definitions, it follows that

$$\|z_T\| \le l a^{\zeta_{m-1}} a^{-\eta_{m-1}} a^{R\overline{\theta}\xi_m} \quad \text{if} \quad T > \tau_m,\ m \ge 2,$$

$$\|z_T\| \le \|\overline{A}(T, s_m) z_{s_m - 1}\| + \|V\| \quad \text{if} \quad T \le \tau_m,\ m \ge 2.$$

In other words,

$$(1 - \mathrm{Ind}\,\{\Omega'\})\,\mathrm{Ind}\,\{\Omega_3\}\|z_T\| \le (1 - \mathrm{Ind}\,\{\Omega'\})\,\mathrm{Ind}\,\{\Omega_3\} l a^{\zeta_{m-1}} a^{-\eta_{m-1}} \beta^{\xi_m},$$

$$\mathrm{Ind}\,\{\Omega'\}\,\mathrm{Ind}\,\{\Omega_3\}\|z_T\| \le \mathrm{Ind}\,\{\Omega'\}\,\mathrm{Ind}\,\{\Omega_3\}\Big(\|\overline{A}(T, s_m) z_{s_m-1}\| + \|V\|\Big).$$

Note that $L < \beta = a^{R\overline{\theta}}$, and if $\mathrm{Ind}\,\{\Omega_3\} \ne 0$ then $m \ge 2$. From (A.13), we have that

$$\mathrm{Ind}\,\{\Omega_3\}\,\|z_{s_m - 1}\| \le \mathrm{Ind}\,\{\Omega_3\} l a^{\zeta_{m-1}} a^{-\eta_{m-1}},$$

$$\mathrm{Ind}\,\{\Omega_3\}\,\|A(T, s_m) z_{s_m - 1}\| \le \mathrm{Ind}\,\{\Omega_3\} l a^{\zeta_{m-1}} a^{-\eta_{m-1}} \beta^{\xi_m}.$$

It means that

$$\mathrm{Ind}\,\{\Omega_3\}\|z_T\| \le \mathrm{Ind}\,\{\Omega_3\}\Big(l a^{\zeta_{m-1}} a^{-\eta_{m-1}} \beta^{\xi_m} + \mathrm{Ind}\,\{\Omega'\}\|V\|\Big).$$

Hence

$$\left(\mathbf{E}\,\mathrm{Ind}\,\{\Omega_3\}\|z_T\|^r\right)^{1/r} \le \mathbf{E}\left(\,\mathrm{Ind}\,\{\Omega_3\}l^r a^{r\zeta_{m-1}}a^{-r\eta_{m-1}}\beta^{r\xi_m}\right)^{1/r} + \mathbf{E}\left(\,\mathrm{Ind}\,\{\Omega_3\cap\Omega'\}\|V\|^r\right)^{1/r}.$$

In other words,

$$\left(\mathbf{E}\,\mathrm{Ind}\,\{\Omega_3\}\|z_T\|^r\right)^{1/r} \le K_1^{1/r} + K_2^{1/r}, \tag{A.17}$$

where

$$K_1 \triangleq l^r\mathbf{E}\,\mathrm{Ind}\,\{\Omega_3\}a^{r\zeta_{m-1}}a^{-r\eta_{m-1}}\beta^{r\xi_m}, \quad K_2 \triangleq \mathbf{E}\,\mathrm{Ind}\,\{\Omega_3\cap\Omega'\}\|V\|^r.$$

By (A.14), it follows that

$$K_2 \le \mathbf{E}\,\mathrm{Ind}\,\{\Omega'\}\|V\|^r \le f(r\rho')^{1/\rho'}\frac{\overline{\delta}(r\rho)^r}{(L-1)^r}. \tag{A.18}$$

By Proposition A.3, the random variable $\mathrm{Ind}\,\{\Omega_3\} = \mathrm{Ind}\,\{m\ge 2\}$ is $\mathcal{F}'$- measurable and $\mathcal{F}''$- measurable; $\mathrm{Ind}\,\{k\ge 1\} = \mathrm{Ind}\,\{m\ge 2\}$. By Lemma A.1 and Lemma A.2, it follows that

$$\begin{aligned}
K_1 &\le \mathbf{E}\,\mathrm{Ind}\,\{\Omega_3\}\|z_T\|^r = \mathbf{E}\,\mathbf{E}\left\{l^r a^{r\zeta_{m-1}}a^{-r\eta_{m-1}}\beta^{r\xi_m}\,\mathrm{Ind}\,\{\Omega_3\}|\mathcal{F}'\right\}\\
&= \mathbf{E}\left(l^r\,\mathrm{Ind}\,\{\Omega_3\}a^{r\zeta_{m-1}}\mathbf{E}\left\{a^{-r\eta_{m-1}}\beta^{r\xi_m}|\mathcal{F}'\right\}\right)\\
&= \mathbf{E}\left(l^r\,\mathrm{Ind}\,\{m\ge 2\}a^{r\zeta_{m-1}}\mathbf{E}\left\{a^{-r\eta_{m-1}}\mathbf{E}\left\{\beta^{r\xi_m}|\mathcal{F}''\right\}\Big|\mathcal{F}'\right\}\right)\\
&\le f(r)l^r\mathbf{E}\left(a^{r\zeta_{m-1}}\,\mathrm{Ind}\,\{k\ge 1\}\mathbf{E}\left\{a^{-r\eta_{m-1}}|\mathcal{F}'\right\}\right) \le l^r f(r).
\end{aligned} \tag{A.19}$$

By (A.12)-(A.19), it follows the inequality (6.2) for the constant (6.9). This completes the proof of Theorems 6.1– 6.3.□

Proof of Theorem 6.4. Let $\overline{l}>0$ be such that (6.5) holds for $l=\overline{l}$. Introduce random events $\Omega_{0,t}\triangleq\{t<s_1\}$, $t=1,...,T$. Let $\varepsilon_t\triangleq 1-\mathbf{P}(\Omega_{0,t})$. We have that $\|z_t\|\le l/\nu$ for $t\le s_1$. Let $C_*=C_*(\rho)$ be the constant defined in Theorem 6.3 with $r=\rho$. By Theorem 6.3, it follows that if $l\ge\overline{l}$, $l/\nu\le 1$, $\nu>La+1$ then

$$\begin{aligned}
\mathbf{E}\|z_t\|^r &\le \tfrac{l^r}{\nu^r}\mathbf{P}(\Omega_{0,t}) + \mathbf{E}(1-\,\mathrm{Ind}\,\{\Omega_{0,t}\})\|z_t\|^r\\
&\le \tfrac{l^r}{\nu^r}\mathbf{P}(\Omega_{0,t}) + \left(\mathbf{E}\|z_t\|^{2r}\right)^{1/2}\left(\mathbf{E}(1-\,\mathrm{Ind}\,\{\Omega_{0,t}\})^2\right)^{1/2}\\
&\le \tfrac{l^r}{\nu^r}\mathbf{P}(\Omega_{0,t}) + C_*(2r)^{1/2}\left(1-\mathbf{P}(\Omega_{0,t})\right)^{1/2} = \tfrac{l^r}{\nu^r}(1-\varepsilon_t) + C_*(2r)^{1/2}\varepsilon_t^{1/2}.
\end{aligned}$$

Let $\kappa\triangleq\min\{i\in\{2,4,...\}:\ i>2r\}$. Similarly the proof of Lemma A.1, we obtain that

$$\varepsilon_t \le \mathbf{P}\left(\exists i\le t: \|\tilde{z}_{i+1}\|>l,\ \|\tilde{z}_i\|\le l\right) \le t\max_{i\le t}\mathbf{P}\left(\|v_i\|>lh\right) \le t\frac{G_\kappa\delta^\kappa}{l^\kappa h^\kappa}.$$

By (6.9), it follows that there exists a constant $c_1>0$ such that $C_*(\rho)\le c_1(l^\rho+1)$ for all $l>\overline{l}$, $\nu>\max(La+1,\overline{l})$, $\rho\in[r,2r]$. Hence

$$\varepsilon_t C_*(2r) \le \frac{TG_\kappa\delta^\kappa}{l^\kappa h^\kappa}c_1(l^{2r}+1)\to 0 \quad \text{as} \quad M\to+\infty, \quad l\to+\infty, \quad \frac{l}{\nu}\to 0.$$

This completes the proof of Theorem 6.4.